2018

The Effect of Increased Quizzing on Retention of Material by Histology Laboratory Students

R. J. Nogrady
Wright State University

THE EFFECT OF INCREASED QUIZZING ON RETENTION OF MATERIAL BY HISTOLOGY LABORATORY STUDENTS

A thesis submitted in partial fulfillment of the
requirements for the degree of
Master of Science

By

R. J. NOGRADY
B.S., Charleston Southern University, 1983

2018
Wright State University

WRIGHT STATE UNIVERSITY

GRADUATE SCHOOL

December 6, 2018

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY R. J. Nogrady ENTITLED The Effect of Increased Quizzing on Retention of Material by Histology Laboratory Students BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science

Barbara E. Hull, Ph.D.
Thesis Director

David L. Goldstein, Ph.D.
Chair, Department of Biological Sciences

Committee on
Final Examination

Barbara E. Hull, Ph.D.

William Romine, Ph.D.

Christa Agiro, Ph.D.

Barry Milligan, Ph.D.
Interim Dean of the Graduate School

ABSTRACT

Nogrady, R. J. M.S. Department of Biological Sciences, Wright State University, 2018. The Effect of Increased Quizzing on Retention of Material by Histology Laboratory Students.

In traditional approaches to teaching, examinations and quizzes have been considered secondary activities, designed to assess and motivate learning, but not generally as teaching tools in themselves (Roediger and Butler, 2011). However, abundant psychological research in laboratory settings indicates that the act of taking a quiz or examination on the material can directly enhance retention of that material in ways which are distinct from and often more effective than restudying of the material. This phenomenon is now referred to as the testing effect. The testing effect hypothesis asserts that (1) repeated retrieval attempts have a longer lasting effect on retention than repeated study attempts, (2) this effect is more pronounced in pure retrieval situations than in recognition situations, and (3) the effect is influenced by the timing of the testing relative to the presentation of the material and the timing of feedback on the results of testing (Wheeler et al., 2003; Karpicke and Roediger, 2008; Toppino and Cohen, 2009; Smith and Kimball, 2010).

Despite the body of clinical research, the applicability of the testing effect to actual educational settings has not been rigorously demonstrated, in part due to the wide range of educational environments and purposes. The primary objective of this project was to investigate the classroom validity of the testing effect in an undergraduate/graduate Vertebrate Histology class. The aim was to determine whether low-stakes testing (non-graded quizzing) could be used to increase students' retention and final exam scores.

A regression discontinuity (RD) design was selected for the research since it offers the treatment hypothesized to be most useful to subjects most in need of such an improved approach (in this case, those with lower test scores). RD is a quasi-experimental design that assigns a cut-off or threshold, with a treatment/intervention allocated to either the above the threshold or the below the threshold group. In this case, the hypothesized superior treatment (non-graded quizzing) was used for students who fared less well on the initial test, while those who scored higher were given an equivalent exposure to same facts in a restudy activity. Results were measured using students' final exam scores.

The experimental results did not provide support for the primary hypothesis; there was no statistically significant improvement in grades for students who underwent non-graded quizzing compared to the restudy group. Potential explanations of this outcome could be researcher error, the complexity of the material to be learned (van Gog and Sweller, 2015) or (in this researcher's opinion) the small sample size (Trochim, 2006).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

Background

While an in-depth discussion of the current understanding of human memory is outside the purview of this paper, a brief discussion is necessary before the testing effect can be explained. Current theory suggests that memory consists of three stages: encoding, storage, and retrieval. Encoding is the beginning of memory and learning and involves exposure to or perception of events. It is by necessity selective. Storage is the retention of information and retrieval is the ability to recover the information from storage when necessary (Anderson, 2000; McDermott and Roediger, 2016; Tulving and Thomson, 1973). These stages of memory formation play an important role in various theories of

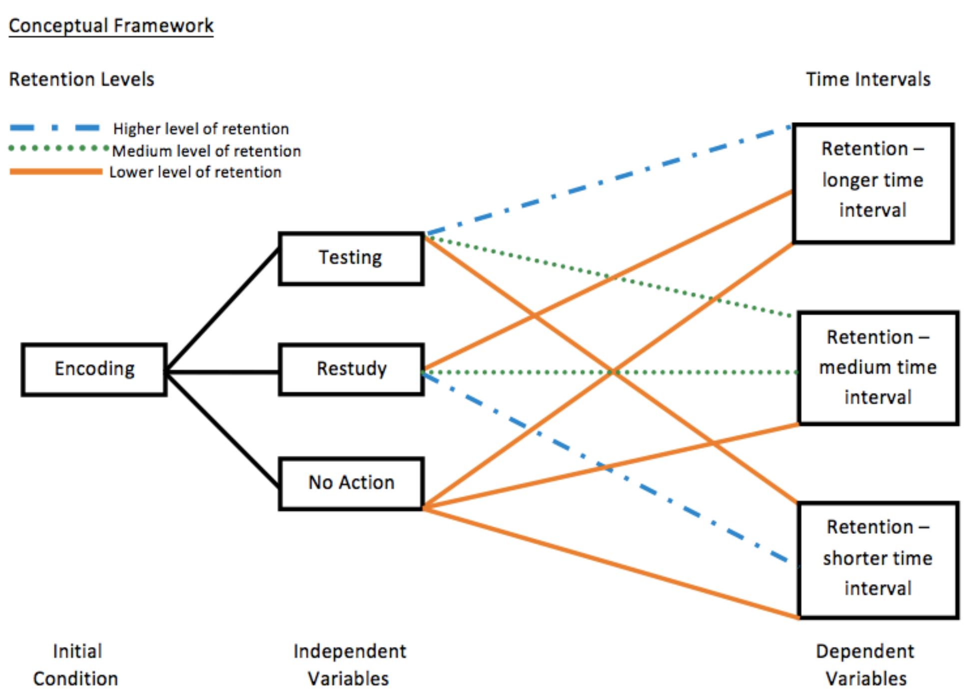


Figure 1. Conceptual Framework for the Testing Effect (Nogrady, 2013)

learning.

The Conceptual Framework model shows the overall idea of the testing effect. On the left side of Figure 1, one primary condition is listed: encoding – the assumption that, before the relative merits of testing or restudy on storage and retrieval of the material can be meaningfully tested, the students must first encode the material.

After completion of the initial encoding or studying, there are three alternatives (the independent variables), two of which (test or restudy) reinforce the initial material studied. The "no action" alternative is included for experimental completeness, even though exposure to the material only once is commonly accepted as not in students' best interest. In practice, “testing” activities can be varied in many ways, such as the timing of the testing relative to initial encoding, provision of feedback, timing of feedback if given, testing of recognition of recall, inclusion of textual and or diagrammatic and spatial information, and maintenance or “dropping out” of material when successfully tested. The common conceptual factor in all of these is exercising retrieval, so these specific types of testing are not included in the diagram (Wheeler, Ewers, and Buonanno, 2003; Karpicke and Roediger, 2008; Toppino and Cohen, 2009; Karpicke and Grimaldi, 2012).

Finally, there are three distinct outcome areas (the dependent variables) for the retention and successful retrieval of material:

1. Longer intervals (greater than two days)
2. Intermediate intervals (less than or equal to 2 days but greater than or equal to 24 hours)
3. Shorter intervals (less than 24 hours)

(Wheeler, Ewers, and Buonanno, 2003; Karpicke and Roediger, 2008; Toppino and Cohen, 2009; Karpicke and Grimaldi, 2012).

This general conception of the testing effect has evolved. Several important studies played roles in the development of aspects of the theory. The influence of this research on the modern understanding of the theory depended on the authors' focus (such as whether their primary interest was the number of “exercise” retrievals necessary for improvement (Myers, 1914), quantifying underlying mental processes or establishing better practical teaching techniques (Jersild, 1929; Spitzer, 1939), and the researchers’ ability to control for outside factors (Keys, 1934)). From the 1910s until WWII, these studies were conducted to optimize application of this effect to actual classroom environments (Jones, 1925; Jersild, 1929; Hertzberg, 1932; Keys,1934) even though the testing effect itself was somewhat loosely defined. After WWII, research on the testing effect decreased and later came to be focused on the specific mental processes involved rather than exploring practical educational effectiveness. Toppino and Cohen (2009) exemplify the “mental processes” stream of literature. The authors maintained that there were issues in preceding studies, one example being experimental design issues in which test groups and control groups differed as test condition participants had dwindling lists of incorrect response items to learn while control participants retained the entire list. Other difficulties included comparability issues between studies using matched pairs of words and others using prose passages for retrieval and testing purposes as well as issues such as how participants encoded material when no instructions were given to the subjects for how that encoding and the restudy were to be done. Toppino and Cohen (2009) asserted that these could be responsible for the "apparent" results of previous studies and that the existence of a testing effect had not been rigorously demonstrated. Toppino and Cohen attempted to remedy these problems by designing a very tightly controlled experiment using cued recall of paired associates and a single, experimenterpaced, initial cued-recall test. Their results verified the testing effect was an actual phenomenon. Wing (2013) provides another example of a similarly "clinical" approach to the testing effect. Wing used neuroimaging to investigate physiological brain states during both the restudy and retrieval processes. He found that the anterior hippocampus, lateral temporal cortices, and medial prefrontal cortex (PFC) responses were stronger with the testing condition than the restudy condition (Wing, 2013). Toppino and Cohen (2009) and Wing (2013) are interesting and shed important light on the fundamental nature of the testing effect--yet neither offers any practical guidelines for more effective classroom use of the phenomenon.

Other non-classroom studies (Clariana, Ross, and Morrison, 1991; Carrier and Pashler, 1992; Cull, 2000; Roediger and Karpicke, 2006b; Butler and Roediger, 2008; Butler, 2010), which try to limit confounding interactions, have helped to rigorously establish and more precisely define the testing effect. This more refined understanding of the testing effect has not been extensively tested in the classroom. Potential outside factors such as class time, number of credit hours taken that semester, family situations, motivation, and other individual circumstances may influence an individual's performance in a classroom; the testing effect may not be useful when some of these factors are present. Some laboratory studies have attempted to simulate stresses upon students in actual classrooms by "pressure manipulations" (actions such as telling research subjects that their score on some activity will determine whether they and a fictitious partner receive a monetary bonus for their participation). Such manipulations may or may not effectively replicate the inherent anxiety associated with taking a class for a grade.

Furthermore, when the design of an experiment is limited to using passages of text or similar material which has no direct association with a particular course of study previously undertaken by the student, the participants are not able to use material learned earlier as scaffolding for later learning. Also, opportunities are typically lacking for the participants in the experiments to use the results of quizzes and other feedback to redefine their level of self-knowledge in a given area and make the appropriate changes in their self-study regimen.

Classroom Applicability

More recent (2011-2015) studies have been conducted in classroom settings (Roediger, Agarwal, McDaniel, and McDermott, (2011); Einstein, Mullet, and Harrison, (2012); Orr and Foster, (2013); Bjork, Little, and Storm, (2014); Khanna, (2015)). Roediger et al. (2011) conducted numerous laboratory studies on the testing effect and theorized that these laboratory studies could be generalized to a classroom. Roediger et al. (2011) implemented a hypothesis (quizzing promotes learning and retention in "wild" environments) in a sixth grade Social Studies class where the teacher was already using quizzing: pretest, post-test, and review (the review was usually two days before each of the four-chapter exams). The three quizzes and the exam had identical questions, only the question order changed. Unfortunately, a majority (74.7% of 162 participants) of surveyed professors stated a preference for giving similar but not identical questions on quizzes and exams (Wooldridge et al., 2014), so the actual classroom environment chosen was not ideally representative. However, this was still a strong early attempt at transitioning modern understanding of the testing effect into actual classroom research. Roediger et al. (2011) found that the quizzed items were better remembered than the read but non-quizzed, lecture-only items.

The authors suggested investigating the effect, where feasible, of spacing rather than massing the quizzing (that is positioning the quizzes apart in time, rather than one after another). They also recommended deliberately designing and inserting a mix of questions such that some would be identical among the quizzes, the final, and verbiage that was taken directly from readings; others would be identical between the quizzes and the final but not verbatim extracts from the readings; and others would be similar in content but differently worded from each other. Comparing student success rates on the different types of questions would address the “teaching the test” concern.

Einstein, Mullet, and Harrison (2012) used an unusual method to demonstrate classroom validity. The authors were concerned with their students’ lack of awareness of the benefits of the testing effect, so they developed class “laboratory” activities designed to inform the students about the testing effect and to demonstrate the testing effect with the students as participants. The authors were trying to determine the classroom validity of the Roediger and Karpicke (2006b) study using two prose passages that had been used in the previous non-classroom laboratory study. Einstein et al. (2012) found that performance was higher in the Study–Test condition than the Study–Study condition, which supported the classroom validity of the testing effect. The authors also provided the compiled classroom results of the Study–Test condition versus the Study–Study condition research to the students so the students could then perform their own analysis. (Einstein, et al., 2012).

Unfortunately, the authors only used the testing effect demonstration in two labs. Further, the prose passages they used were reading comprehension paragraphs from the Test of English as a Foreign Language (Roediger and Karpicke, 2006b), which covered topics that were not applicable to the psychology course in which the experiment was performed. Even though the researchers conducted the study in a classroom environment, they apparently modified elements of the teaching and testing processes to more closely resemble laboratory conditions. Continuing the semester using within-subjects design would have given the students more reinforcement for the benefits of integrating testing into their studies. Since the authors were interested in encouraging students to use self-testing in their studies, future research should include instruction on methods of doing that. Examples might include study guide questions; flashcards; read, write and say. These materials could be used to highlight the effectiveness of self-testing in addition to exposure to the ideas of spacing versus massing and interleaving or layering the different class topics instead of blocking (studying one subject until completed).

Orr and Foster (2013) reiterate Bjork and Bjork's (2011) premise that for students to achieve success on college classroom exams, they must store and integrate the required material, then successfully retrieve the stored information; and therefore, low-stakes frequent quizzing should benefit students as it allows them to practice retrieval and strengthen the retrieval effort (Bjork and Bjork, 2011). Orr and Foster (2013) aimed to further validate Bjork and Bjork's (2011) work by requiring students to complete online quizzes before each exam, which was hypothesized to then positively impact students' exam scores. Orr and Foster (2013) found that the students in the "100% quiz takers" category outperformed the 0% quiz takers for each group. The authors did not, however, show any results from a control group undergoing a rereading treatment, which required a similar amount of time as the quizzing, so the quiz-taking group's higher grades may just have been due to increased exposure to the material. Also, the self-selected 0% quiz takers may have other issues; if so, even if this group had taken 100% of the quizzes, it may still not have performed well (Orr and Foster, 2013).

Bjork, Little, and Storm (2014) point out that testing performance is not always the same thing as learning (Blodgett, 1929; Postman and Tuma, 1954; Adams and Reynolds, 1954; Stelmach, 1969; as cited in Bjork, Little, and Storm, 2014), either from the student's subjective perspective or the instructor's sometimes short-term perspective (Bjork, 1994). Students may believe they have fully mastered material by reiterating the material immediately after mass study, yet a week later, be unable to retrieve the facts that came so quickly from cramming (Roediger and Karpicke, 2006b). Bjork, Little, and Storm (2014) then reference Bjork's 1994 work which combined theoretical results of some previous empirical studies discussing what Bjork referred to as "desirable difficulties" (seeming impediments to study which improve longer-term retention). These "desirable difficulties" consist of:

1) Distributed practice (spreading out studying versus back-to-back or "massing" study) (Dempster, 1988; Glenberg, 1992; Lee and Genovese, 1988; as cited in Bjork, 1994)

2) Changing practice conditions (for example, changing locations – different areas of the home or different buildings; or varying types of study – flashcards, read, write and say, or instructing a classmate) (Shea and Morgan 1979; Reder, Charney and Morgan 1986; as cited in Bjork 1994; Smith, Glenberg, and Bjork 1978 as cited in Bjork and Bjork 2011)

3) Contextual interference (interspersing different topics – studying chemistry in between histology sessions) (Battig, 1979; Shea and Morgan 1979; Mannes and Kintsch, 1987; as cited in Bjork, 1994; Rohrer and Taylor 2007 as cited in Bjork and Bjork 2011)

4) Testing (practicing retrieval reinforces the mental link plus gives one experience performing what is required on an exam) (Izawa, 1970; Landauer and Bjork, 1978; Rea and Modigliani, 1985; as cited in Bjork, 1994).

Bjork (1994) concluded that these difficulties may lower immediate performance yet lead to enhanced long-term learning.

Bjork, Little, and Storm (2014) had two goals: 1) to investigate the classroom validity of the testing effect and 2) to explore the potential classroom effects of retrieval-induced forgetting (a phenomenon seen in the laboratory in which exercised recall of some information elements in a studied set can lead to impaired access to related, non-target information elements) (Bjork, Little, and Storm, 2014). They used a very cleverly constructed bank of questions applied to various treatment groups, which allowed them to compare retention, transfer and other outcomes within one test. Bjork, Little, and Storm's (2014) results supported the idea that the testing effect can be generalized into actual classroom settings. The students' performance increased for questions that tested the same concept quizzed earlier, but with slightly different word choice, which supported the hypothesis that there is no substantial retrieval-induced forgetting with multiple-answer quizzes. (Bjork, Little, and Storm, 2014).

Bjork, Little, and Storm (2014) indicate they believe the most influential factor for the students' increased performance was the quizzing. However, they could not rule out a metacognitive effect with the quizzing allowing students to determine what areas they were deficient in and increase their studies in these areas. Also, some of the benefit may have come from spacing since the quizzes were given several days after the associated lecture. (Bjork, Little, and Storm, 2014). Future studies attempting to control for a metacognitive effect would seem to require more extensive changes to the overall course curriculum, potentially introducing more variables than are eliminated. However, varying the spacing of the quizzes using a within-group design could determine whether spacing impacted the results.

Khanna (2015) postulated that anxiety potentially affected the students' retention in high stakes quizzing regimens in actual classrooms, so the study investigated research with three conditions: no quizzing, ungraded quizzing (low-stakes) and graded quizzing (high-stakes). The quiz and exam questions were constructed to provide a mix of items at the knowledge, comprehension, and application levels (Bloom, Engelhart, Furst, Hill, and Krathwohl (1956) as cited in Khanna (2015)). The author found that both quizzed groups outperformed the unquizzed group. Further, graded quizzes were more stressful and also resulted in better attendance than ungraded quizzes. However, the graded quizzes did not produce better exam grades (Khanna, 2015).

All of these studies used the randomized control trial (RCT) experimental design. In many situations, it is not feasible to use the same group of people for both the program and control; therefore, the RCT is used to create groups that are probabilistically equivalent. Randomization of assignment to control or treatment groups ensures the groups are statistically similar so that any post-treatment differences can be attributed to the treatment used versus any pre-existing differences between the groups (Trochim, 2006).

Nonetheless, the RCT design may not always be ideal. According to Emanuel, Wendler, and Grady (2000), consideration should be given to selecting subjects who will derive the most benefit from a proposed treatment, but the use of RCT for a testing effect study ensures that both treatment and control groups will have a mix of both stronger and weaker students. Further, a student who has strong study habits and consistently receives high grades may benefit less from a particular treatment than a student who has average grades; thus, an RCT design could substantially understate the effectiveness of the treatment. As Roediger et al. (2011) indicated while discussing their study's limitations, their effect sizes were small (around 10%). However, their non-tested (control group) performance results were high (for example, the non-tested performance mean for one experiment was 81%), which may have artificially limited or capped the potential improvement range for tested students to the remaining 19%, and thereby underreported the strength of the effect (Roediger et al., 2011).

One approach to dealing with these concerns is the Regression Discontinuity (RD) Design. Trochim (2006) defines RD as a "pretest-posttest program-comparison group strategy" which uses a threshold value to determine assignment to a treatment (there is no requirement for the premeasure and the post measure to be the same). In discussing both RCTs and RD Designs, Trochim comments that the main advantage of RD design is that it is applicable for targeting a treatment to those who most need it when the traditional experimental designs are not feasible or are ethically questionable (Trochim, 2006). Furthermore, the RD design lends itself to situations in education, like the impact of delayed entry into kindergarten and the impact of class size reduction (Jacob, Zhu, Somers and Bloom, 2012). For RD analysis to be successful, there cannot be a discontinuity in the distribution of preprogram (before treatment) test results, since the assumption is that the observations just below the cut-off are similar to those just above the cut-off. After a treatment, the comparisons are made between the observations that are relatively close to the cut-off and any differences at the threshold are defined as the effect size (Moscoe, 2015). To maximize the potential benefit for the students involved in the study, the Regression Discontinuity (RD) Design was chosen for this research.

II. METHODS

Participants

The Regression Discontinuity Design was used in three upper level, registered sections (n=14, 13, and 7) of a vertebrate histology course offered at a Midwestern university. Six additional students in an unregistered seminar also participated, for experimental completeness, producing a total of forty subjects. There were 28 females and 12 males.

Structure

Vertebrate Histology Laboratory semesters were 14 weeks long. All students completed an initial assessment, two block exams and a final comprehensive exam (within which the first assessment was incorporated). The registered students received 25 in-class lectures and 25 in-lab review lectures with no lab reviews given the days of the exams. Up until the first exam, all 34 students in registered sections were treated equally, and this first exam was used as the RDD preprogram test (22 students above the cutoff of the mean of 93, who were later given summary readings, and 12 students below the cutoff, who were later given low-stakes learning exercises). Students were required to annotate their weekly time spent studying histology and these times were averaged for the semester for each student and used as a covariate in the analysis.

For experimental completeness, there was a weekly seminar for non-registered volunteer students. These participants represent the “no action” block in the conceptual diagram, Figure1. These six students were given incentives to attend regularly. Their exams were not graded for credit, and they were not part of the RDD analysis since they only received an initial encoding. These participants were only exposed to the material once, via class lectures, and they received a lecture after each of the two block exams.

Materials.

Lab Books. Registered students were required to maintain a composition book, which included their hand drawings of the idealized versions of the tissues they viewed through the microscope (facts about these cells and tissues were lectured on and were covered in required readings). These drawings are designed to help internalization of the nuances of the structures and identification of features of individual cells and their related tissues and were intended to help the students recognize these small but specific details in the micrographs during their learning exercises, summaries, and the exams. The unregistered students did not produce these drawings.

Learning Exercises. Since the focus of this study was the effects of increased quizzing in an upper level histology class, the treatment—increased quizzing—was used at every opportunity, specifically in every laboratory lecture session except those in which there was a lab exam. These quizzes (which were presented to students as "low-stakes learning exercises" to reduce anxiety over the term "quiz") were adapted from for-credit quizzes used two years prior in the same class. The quizzes were created by laboratory teaching assistants (who had previously taken the course) with the specific intention of covering areas that were typically troublesome for many students. The quizzes were low-stakes in that performance on the quizzes (number of correct answers) had no impact on class grades but participating in the quizzing process (or the control group process) allowed the students to receive attendance credit. The learning exercises were only available online for the last 15 minutes of each laboratory session (except those containing a for-credit, traditional lab exam). The quizzes were digitized; students received a code to access the correct treatment in Pilot, the university's learning management system. The quizzes consisted of a micrograph and one or a series of short-answer questions associated with the topic (see Appendix F for a sample question on Pilot, the university's computerized learning management system / class electronic bulletin board). Pilot was set up so that the students undergoing the quizzing treatment had to submit an answer before they could proceed to the feedback. Since the students' wording of their answers could vary widely, both the question and the feedback included an instructor-created answer block in which the students would place an "x." That x was the only quiz response that the computer recognized, and it was used only to determine attendance credit.

The control group and the treatment group were presented with the same micrographs as part of each laboratory session; however, the control group's meetings were accompanied with a written summary they were to read during the last 15 minutes whereas the treatment group was given a quiz with an average of ten questions which covered the same information as the summary. Immediate feedback on quiz results was provided. The unregistered students did not receive summaries or quizzes.

Exams. The first two exams were worth 50 points and the final exam was worth 60 points; all three consisted of micrographs with related questions. None of the question/micrograph combinations were identical to those on the learning exercises.

Procedures

A combined-class lecture section met twice a week in 80-minute lecture periods. The three separate lab sections met twice a week in 110-minute periods. Throughout the semester, all three lab sections covered the same 16 chapters in the same order. For approximately the first 30 minutes of the lab, the lab instructor lectured on some of the more difficult material covered in the combined-class lecture and answered questions. The students then worked with their microscopes and on their drawings. There was generally one teaching assistant for every seven students to answer any questions on microscope viewings or drawings. In the final fifteen minutes of each lab section (except exam days), the students logged in to separately coded areas for treatment and control and participated in the appropriate treatment. The students below the threshold were given the low-stakes computerized quiz treatment while the students above the threshold were given a digital summary reading.

III. RESULTS

Descriptive statistics are presented in Table 1. The first exam consisted of questions that were identical (but for minor editing) to those given to the previous two histology classes on their first exam; those exams were not released to students. The mean was chosen to serve as the threshold for dividing treatment and control groups before the presentation of the first exam / Preprogram test; based on the previous two classes (2015 exam1 mean, 88%; 2016 exam1 mean 84.5%; rounded to the nearest half point). The Preprogram test mean was expected to be not higher than 88%, however, in the event, the mean score was just over 93%. Although it was recognized that this high mean (and using it as the treatment threshold) might make it challenging to establish the statistical significance of any observed effects, it was decided that this was better than introducing an arbitrary cutoff. Students that scored lower than the mean were placed in the treatment group (low-stakes quizzing; n=12) and students that scored higher than 93 were placed in the control group (summary readings; n=22). Note that the exam 1 means are higher than the final exam means. Lower final exam scores are typical in this class (leading to a requirement that students' score at least a B on the final to earn an A for the course). The average for all sections on the final exam was considerably lower than that for exam 1. However, the treatment group experienced a substantially smaller drop than did the control group, which provides some support to the hypothesis—but this result could have been the result of pure chance, given the relatively small samples and high variability of test scores.

Table 1. Descriptive data by group

Group	n	Test	Mean	Standard Deviation	Max	Min
All	34	Exam1/ Preprogram test	93.28	7.49	103.50	72.00
		Final/ Post-program test	88.65	10.68	103.75	52.50
Control	22	Exam1/ Preprogram test	97.77	9.99	103.50	93.50
		Final/ Post-program test	91.23	8.25	103.75	72.92
Treatment	12	Exam1/ Preprogram test	85.04	5.89	92.50	72.00
		Final/ Post-program test	83.92	13.22	100.00	52.50

Cronbach's alpha (CA) was run (SPSS 25) to determine the internal consistency of the post-treatment data. All questions that were answered correctly by all students were removed before CA was executed and the results are listed in Table 2.

Table 2. Cronbach's Alpha on the Post-treatment Exam

Cronbach’s Alpha	Cronbach’s Alpha Based on Standardized Items	N of Items
.918	.921	110

Since the CA for the post-measure exam was higher than the recommended 0.6, several univariate analyses were run on the data. None of the SPSS outcomes resulted in significant p values (see Appendix A) as expected. Since the initial premeasure exam mean was higher than expected and also, by design, the pretreatment exam was different from the post-treatment exam, factor analysis/principal components was employed. Factor analysis/principal components are used to find the fewest factors (observed variables that have similar response patterns) that explain the most variance in the original data.

Factor Analysis. (Principal Components Method).

The experimental data were reduced into components (sometimes called factors), in order to investigate the total variability associated with particular variables and groupings of variables included in the analysis. The selection criteria for a principal component was an eigenvalue larger than one (see Table 1 for excerpt, see Appendix C for all the associated eigenvalues).

Table 3. Eigenvalues above 1 - Components 1 through 7; Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings a
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	15.983	14.53	14.53	15.983	14.53	14.53	9.358
2	9.736	8.851	23.381	9.736	8.851	23.381	7.698
3	7.977	7.252	30.633	7.977	7.252	30.633	7.181
4	7.617	6.925	37.558	7.617	6.925	37.558	7.67
5	5.675	5.159	42.717	5.675	5.159	42.717	4.836
6	5.301	4.819	47.536	5.301	4.819	47.536	4.808
7	5.074	4.613	52.149	5.074	4.613	52.149	4.237
8	4.827	4.388	56.536	4.827	4.388	56.536	4.937

Extraction Method: Principal Component Analysis.
a When components are correlated,
sums of squared loadings cannot be added to obtain a total variance.

The initial eigenvalues are the variances of the components. Only eight of the components are displayed in Table 1. These eight account for over half (56%) of the variance seen in the data, while the remaining 18 components (9 through 26) account for 44% of the variance. The extraction sums of squared loadings rows will correspond to the number of components retained, in this case 26 components. The rotation sums of squared loadings represent the variance after the rotation.

The pattern matrix used for component identification is located in Appendix D. The specific final examination questions with factor loadings at or above 0.3 comprising that component were examined as a group to identify any linguistic or histological characteristics common to those questions but absent or less pronounced in other questions. (see Appendix E for an example - PCA column 2). This review of characteristics was accomplished both by the researcher and by an experienced histology professor. The latter determined that Column 2 PCA, which contained twelve final exam questions with a Cronbach's alpha 0.857, were in fact similar in that they require the student to integrate knowledge of staining processes and functional recognition to a degree greater than that seen in other questions, and the twelve questions with a Cronbach's alpha 0.799 in Column 5 PCA require the ability to deductively apply knowledge of general tissue organization rules and concepts to isolate and identify significant unique features of a particular tissue sample more than other test questions do. These were the only two components that showed a consistent significance.

A significance of our treatment variable within the PCA indicates that the treatment students performed better than the control students as projected from the first exam. Although this supports the hypothesis, this evidence is not conclusive. It is also possible that the learning exercises did not improve treatment group students' retention or recall of material per se, but did impress upon them that successful tissue identification requires both recall/recognition of previous material and evidence-based reasoning about the specimen at hand, while control students may have been encouraged to believe that simple recognition alone was sufficient.

When factor analysis was performed on the post-treatment exam (minus the "all correct" answers), 22 positive factors were disclosed in the Pattern Matrix. Only questions with a factor loading greater than .3 and an eigenvalue greater than one were considered for univariate analysis. Each individual factor (which consisted of a grouping of eight to sixteen final exam questions) was examined for similarities between the questions, to include: tissue types, necessity of recognition of a cell's internal characteristics, shapes of cells, orientation of micrographs, contrasts as a result of staining, and question “linguistic” variations such as question length and/or complexity. The factors with more than seven questions were converted into dummy variables representing correct, partially correct and incorrect answers. These nine factors were then run through Cronbach’s alpha with all scoring above the .6 requirement. The factors were then reconstructed into scores for each student for each factor, and univariate analysis was performed on each factor with student lab attendance as a covariant. See Appendix B.

Only limited significance was found on three factors. Model 2 (fixed factor – treatment category, covariates – precut (pretreatment exam) and attendance average) was significant in factor 2 [$F(1,30)=5.722$, $p=0.023$], factor 5 [$F(1,30)=4.288$, $p=0.047$], and factor 22 [$F(1,30)=4.294$, $p=0.047$]. Although some models were technically significant, both the Bonferroni effect and the fact that no common characteristics were discerned within the factors makes this weak evidence for any actual effect.

Since column width was limited on the tables the following abbreviations were used:

Dep. Var. – Dependent Variables

Sig. – Significance

Adj. – Adjusted

<u>Fixed Factor Used</u>

P3 – Final exam

<u>Covariates Used</u>

PreCut – Pretreatment Exam1 with the mean of 93 subtracted

PreCutSqrd – PreCut squared

PreCutCubed – PreCut cubed

C0 T1 – Treatment Category

C0 – Control

C1 – Treatment

StudyTm – The self-reported amount of study per week averaged for the semester

Attend. – Attendance for the semester

Interact1 – Interaction 1; Precut times the treatment variable

Interact2 – Interaction 2; Precut Squared times the treatment variable

Interact3 – Interaction 3; Precut cubed times the treatment variable

FA – Principal Component (Factor Analysis)

<u>Analyses Used</u>

Anal 1- Analysis 1; which included PreCut, AvgStudyTm and Interact1

Anal 2- Analysis 2; which included PreCut, AvgStudyTm without Interact1

Anal 3- Analysis 3; which included PreCut, without AvgStudyTm, with Interact1

Anal 4- Analysis 4; which included PreCut, without AvgStudyTm, without Interact1

Anal 5- Analysis 5; which included PreCutSqrd, PreCut, AvgStudyTm, Interact1, and Interact2

Anal 6a- Analysis 6a; which included PreCutSqrd, PreCut, AvgStudyTm, Interact1, and without Interact2

Anal 6b- Analysis 6b; which included PreCutSqrd, PreCut, AvgStudyTm, without Interact1, and with Interact2

Anal 6c- Analysis 6c; which included PreCutSqrd, PreCut, AvgStudyTm, without

Interact1, and without Interact2

Anal 7- Analysis 7; which included PreCutSqrd, PreCut, without AvgStudyTm, with Interact1, and with Interact2

Anal 8- Analysis 8; which included PreCutSqrd, PreCut, without AvgStudyTm, without Interact1, and without Interact2

Anal 9- Analysis 9; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, Interact1, Interact2, and Interact3

Anal 10a- Analysis 10a; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, Interact1, Interact2, and without Interact3

Anal 10b- Analysis 10b; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, Interact1, without Interact2, and with Interact3

Anal 10c- Analysis 10c; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, without Interact1, with Interact2, and with Interact3

Anal 10d- Analysis 10d; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, Interact1, without Interact2, and without Interact3

Anal 10e- Analysis 10e; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, without Interact1, without Interact2, and with Interact3

Anal 10f- Analysis 10f; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, without Interact1, with Interact2, and without Interact3

Anal 10g- Analysis 10g; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, without Interact1, without Interact2, and without Interact3

Anal 11- Analysis 11; which included PreCutCubed, PreCutSqrd, PreCut, without AvgStudyTm, with Interact1, Interact2, and Interact3

Anal 12- Analysis 12; which included PreCutCubed, PreCutSqrd, PreCut, without

AvgStudyTm, without Interact1, without Interact2, and without Interact3

The same preceding analyses were used for each of the eight largest components: FA2, FA3, FA5, FA10, FA11, FA12, FA22 and FA25.

IV. CONCLUSIONS

Although lack of significance meant that this study had no theoretical implications, in practical terms, this research proved quite productive for both focusing research methods in this particular environment and improving classroom teaching techniques. The major problem with this study was the small sample size. The research was refined over four years, correcting potential confounding issues and improving the design. These changes meant that the research from each year could not be combined to form a larger sample size. The first year was a pilot investigation with the researcher and two additional graduate teaching assistants (GTAs) each managing a lab. This design used non-computerized non-graded quizzes and a lecture summary review, which necessitated that the control and treatment participants be in separate lab sections. The researcher made minimal input to the other two GTA's labs; therefore the lectures were not fully standardized. Initial data collection occurred the second year of research. A randomized control trial (RCT) design, following the same format as the pilot, was used. However, the lectures were standardized with all lab instructors using the same PowerPoint slides. Since there could be issues with different lecture styles between the teaching assistants, a computerized version was developed for the third year (second year's data collection). This digitizing also helped by removing the potential for differing styles of instruction from effecting treatment and control groups differently since each lab (and each instructor) would have both control and treatment participants. Based on considerations identified by Emanuel, Wendler, and Grady (2000), there was a question on the effectiveness of this procedure since the RCT included students who potentially would not benefit from the hypothesized better treatment. Therefore, a regression discontinuity (RD) design was implemented for the fourth year (third data collection period). Even though RD was a more suitable design on theoretical grounds, it was not optimal regarding the ability to demonstrate statistical significance within the given small sample size. The RD design works best with at least three times the sample size as RCT (Trochim, 2006). The RCT design could have been used for the third data collection period thereby increasing the sample size to include both the second and third data collection periods.

Also, the investigation could have been more comprehensive as far as distributed practice (spacing between encoding and quizzing) and changing the practice conditions (changing locations or time of day). The distributed practice could have been improved by the spacing of the lecture and low-stakes quizzes (ungraded learning exercises) at least two days apart, rather than the same day as the lecture. Furthermore, both the practice conditions and distributed practice could have been improved by coordinating the ungraded learning exercises so that identical questions could be asked two weeks later in either a lecture (different classroom and spacing) or lab (spacing) graded quizzes. Furthermore, classification of questions by Bloom's taxonomy could have been used to determine whether the testing effect would be more useful with material at the lower levels (knowledge and comprehension) or the higher levels (analysis and synthesis).

As a result of this research, the histology labs are being restructured to enhance student learning by incorporating lessons learned during the conduct of the investigation. Furthermore, other actions will be taken to contribute to further changes to the labs.

Even though the undergraduate teaching assistants were not part of the study, they showed much more awareness of the material from having to explain/teach their peers.

Therefore, an interteaching (in which students form small groups within the class, discuss the homework previously assigned and complete their lab books within the lab) approach could be used in future labs. Even though the learning exercises will no longer be utilized, on the graded quizzes, students could have the option to correct their mistakes or explain their reasoning on how they arrived at the wrong answer for half-credit. This process should take the quizzes from the high-stakes to the low-stakes level.

Electronic polling “clickers” are convenient and can provide anonymous interaction and rapid feedback to students, and therefore could be an effective means for implementing quizzing. The clickers could allow distributed spacing within a lecture instead of massed spacing at the end of the lecture plus have the added feature of immediate feedback in addition to awareness of how an individual’s knowledge of an item compares to the class as a whole. Clickers are being investigated for possible incorporation into the lab in the future.

As another option, the quizzing could be embedded and dispersed in the post small group follow-up lecture. This lecture placement provides for spacing between the introduction of the material in the previous lab and quizzing and also provides for removing the massing of all quiz questions together.

Other changes to the lab structure could be made to address students' perceptions about particular teaching methods. For example, in all four years of increased quizzing most students showed great initial resistance to quizzing. However, by the end of the semester, many students made unsolicited comments that the quizzing had been very beneficial, helping them not only remember specific items more effectively but also increasing their awareness of which topics that they thought they had fully mastered but required more work. Some students who did not come to that realization may not have fully experienced the potential benefits of quizzing. Furthermore, students made unsolicited comments on how the “wet labs” of sectioning and staining contributed to their overall understanding of histology. Therefore, a survey or other method of querying the students on their perception of quizzing and other teaching techniques and how quizzing or other techniques could be used to enhance the classroom experience will be added. This survey could be given at the end of the first week of quizzing and the end of the semester. Additionally, now that the study is finished, information about the testing effect can be incorporated into the labs so the students could be made aware of the potential gains from quizzing.

Volunteers could be re-examined during the summer or fall to retake a mixture of the exam questions to determine the amount of forgetting that takes place and to determine which specific topics are best remembered and which are forgotten. This data could then be used to modify future handling of these topics.

Finally, even though this information have been mentioned earlier, these points are important enough to be addressed again, together. Bjork and Bjork (2011) addresses several “desirable difficulties,” learning techniques that may be counter-intuitive but nonetheless appear to help students learn much more effectively. These include distributed practice, changing practice conditions, contextual interference, and testing. Efforts could be made by instructors at all levels to both inform students of how beneficial these practices could be for retention of material over time, and just as importantly, review instructor-designed in-class activity and homework assignments to ensure they facilitate use of these techniques. Specifically, students could be informed that:

-rather than repeatedly studying in the same location, it can be more effective if they use several study locations. This is part of “Varying the Conditions of Practice.”

-rather than “massing” study of a particular topic, spacing or distributed practice can lead to better long-term retention and recall. Despite apparent rapid gains in mastery, in which students can correctly recall large amounts of material during or immediately after a “massed” study session, learning theory suggests that memory strength should not be confused with retrieval strength. The immediate repetitions used in massing or “cramming” do not require accessing long term memory but are accessing working memory, and so the material may not be readily retrievable during a test or as a scaffolding for later learning.

-alternating (interleaving) study topics is a relatively simple strategy that can be used to achieve the benefits of spacing and contextual interference.

Other applications of these ideas would require instructors to consider modifying aspects of their course delivery. For example, students may complain if lectures and labs do not align with the same schedule of topics, but Bjork and Bjork (2011) indicate that it is better for long-term retention if the topics in the lab immediately after lecture are not the same as the lecture. Histology lab lecture / lab sequencing will be reviewed to see where such opportunities can be incorporated.

As the focus of this research was on the testing effect, testing activity in the histology class will also be closely reviewed for applicability of Bjork and Bjork’s recommendations, with special emphasis on creating opportunities for students to combine quizzing-based retrieval exercises with spacing and interleaving, use of quizzing throughout a lecture instead of massing at the end of the lecture, and quizzing multiple topics within one quiz to achieve interleaving. Additionally, methods to encourage or facilitate student self-quizzing activity will be explored.

V. REFERENCES

Anderson, J. R. (2000). Learning and memory: An integrated approach (2nd ed.). Hoboken, NJ, US: John Wiley and Sons Inc.

Adams, J. A., and Reynolds, B. (1954). Effect of shift in distribution of practice conditions following interpolated rest. *Journal of Experimental Psychology*, 47, 32–36.

Battig, W. F. (1979). The flexibility of human memory. In L. S. Cermak and F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 23 – 44). Hillsdale, NJ: Lawrence Erlbaum Associates.

Bjork E.L. and Bjork R.A. (2011). Making things hard on yourself, but in a good way: creating desirable difficulties to enhance learning. In: Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society, ed. MA Gernsbacher, RW Pew, LM Hough, and JR Pomerantz, New York: Worth, 56–64.

Bjork, E. L., Little, J. L., and Storm, B. C. (2014). Multiple-choice testing as a desirable difficulty in the classroom. *Journal Of Applied Research In Memory And Cognition*, 3(Cognition and Education), 165-170. doi:10.1016/j.jarmac.2014.03.002

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe, and A. Shimamura (Eds.), Metacognition: Knowing about knowing (pp. 185–205). Cambridge: MA: MIT Press.

Blodgett, H. C. (1929). The effect of the introduction of reward upon the maze performance of rats. *University of California Publications in Psychology*, 4, 113–134.

Bloom, B. S. (1966). Twenty-Five Years of Educational Research. *American Educational Research Journal*, 3(3), 211.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., and Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of goals. Handbook I: Cognitive domain. New York, NY: Longmans, Green.

Butler, A.C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Educational Psychology*: *Learning, Memory and Cognition*, 36(5), 1118-1133.

Butler, A.C. and Roediger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory and Cognition*, 36(3), 604-616.

Carrier, M. and Pashler, H. (1992). The influence of retrieval on retention. *Memory and Cognition*, 20, 633-642.

Clariana, R. B., Ross, S. M., Morrison, G. R. (1991). The Effects of Different Feedback Strategies Using Computer-Administered Multiple-Choice Questions as Instruction. *Educational Technology Research And Development*, (2), 5.

Cull, W. L. (2000). Untangling the benefits of multiple study opportunities and repeated testing for cued recall. *Applied Cognitive Psychology*, 14(3), 215-235.

Dempster, F. N. (1988). The Spacing Effect: A Case Study in the Failure to Apply the Results of Psychological Research. *American Psychologist*, *43*(8), 627–634.

Einstein, G. O., Mullet, H. G., and Harrison, T. L. (2012). The Testing Effect: Illustrating a Fundamental Concept and Changing Study Strategies. *Teaching Of Psychology*, 39(3), 190-193.

Emanuel, E.J., Wendler, D., Grady, C., (2000) "What Makes Clinical Research Ethical?" *JAMA* 283, no. 20 2701-2711.

Glenberg, A. M. (1992). Distributed practice effects. In L. R. Squire (Ed.), Encyclopedia of learning and memory. (pp. 138-142) New York, NY: Macmillan Publishing Co, Inc.

Hertzberg, O. E., Heilman, J. D., and Leuenberger, H. W. (1932). The value of objective tests as teaching devices in educational psychology classes. *Journal of Educational Psychology*, 23(5), 371-380. doi:10.1037/H0072896

Izawa, C. (1970).Optimal potentiating effects and forgetting-prevention effects of tests in paired-associate learning. *Journal of Experimental Psychology*, 83, 340–344.

Jacob, R., Zhu, P., Somers, M.A., and Bloom, H. (2012). A practical guide to regression discontinuity. MDCR. Accessed 21 Oct 2016, from http://citeseerx.ist.psu.edu/viewdoc/download?Doi=10.1.1.671.4723&rep=rep1&type=pdf

Jersild, A. T. (1929). Examination as an aid to learning. *Journal of Educational Psychology*, 20(8), 602-609. Doi:10.1037/H0070530

Jones, H. E. (1923). Experimental Studies of College Teaching. *Archives of Psychology,* No. 68, 1923.

Karpicke, J. D., and Grimaldi, P. J. (2012). Retrieval-Based Learning: A Perspective for Enhancing Meaningful Learning. *Educational Psychology Review*, (3), 401.

Karpicke, J. D. and Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, 319(5865), 966-968.

Keys, N. (1934). The influence on learning and retention of weekly as opposed to monthly tests. *Journal of Educational Psychology*, 25(6), 427-436. doi:10.1037/H0074468

Khanna, M. M. (2015). Ungraded Pop Quizzes: Test-Enhanced Learning without All the Anxiety. *Teaching Of Psychology*, 42(2), 174-178.

Landauer, T.K., and Bjork, R.A. (1978). Optimum rehearsal patterns and name learning. In M.M. Gruenberg, P.E. Morris, and R.N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). New York: Academic Press.

Lee, T. D., and Genovese, E. D. (1988). Distribution of practice in motor skill acquisition: Learning and performance effects reconsidered. *Research Quarterly for Exercise and Sport*, *59*, 277-287.

Mannes, S. M., and Kintsch, W. (1987). Knowledge organization and text organization. *Cognition and Instruction*, 4, 91–115.

McDermott, K., and Roediger, H. (2016). Memory (Encoding, Storage, Retrieval). *Noba textbook series: Psychology.* Champaign, IL*:* DEF publishers. doi: https://doi. org/noba project. com.

Moscoe, E., Bor, J., and Bärnighausen, T. (2015). Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice. *Journal of Clinical Epidemiology*, 68(2), 132-143

Myers, G. C. (1914). Recall in relation to retention. *Journal of Educational Psychology*, 5(3), 119-130. Doi:10.1037/H0075769

Orr, R., and Foster, S. (2013). Increasing Student Success Using Online Quizzing in Introductory (Majors) Biology. *CBE - Life Sciences Education*, *12*(3), 509-514.

Postman, L., and Tuma, A. H. (1954). Latent learning in human subject. *American Journal of Psychology*, 67, 119–123.

Rea, C. P., and Modigliani, V. (1985). The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning: Journal of Practical Research and Applications, 4*(1), 11-18.

Reder, L.M., Charney, D.H., and Morgan, K.I. (1986). The role of elaborations in learning a skill from an instructional text. *Memory and Cognition*, 14, 64–78.

Roediger, H. L., III, Agarwal, P. K., McDaniel, M. A., and McDermott, K. B. (2011). Test-Enhanced Learning in the Classroom: Long-Term Improvements from Quizzing. *Journal Of Experimental Psychology: Applied* 17, no. 4: 382-395.

Roediger, H. L., III., and Karpicke, J. D. (2006b). Test-enhanced learning: Taking tests improves long-term retention. *Psychological Science, 17,* 249 –255.

Roediger, H. L., and Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20-27.

Rohrer, D. and Taylor, K. (2007). The shuffling of mathematics practice problems improves learning. *Instructional Science*, 35, 481-498.

Shea, J.B. and Morgan, J.B. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *J Exp Psychol: Hum Learn Mem*. 5:179–187.

Smith, S. M., Glenberg, A. M., and Bjork, R. A. (1978). Environmental context and human memory. *Memory and Cognition*, 6, 342-353.

Smith, T. A., and Kimball, D. R. (2010). Learning from Feedback: Spacing and the Delay-Retention Effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 80-95. Doi:10.1037/a0017407

Spitzer, H.F. (1939). Studies in retention. *Journal of Educational Psychology*, 30, 641-656.

Stelmach, G. E. (1969). Efficiency of motor learning as a function of intertrial rest. *Research Quarterly*, 40, 198–202.

Toppino, T. C. and Cohen, M. S. (2009). The testing effect and the retention interval: questions and answers. *Experimental Psychology*, 56(4), 252-257.

Trochim, W. M. (2006). The Research Methods Knowledge Base, 2nd Edition. Internet WWW page, at URL: <http://www.socialresearchmethods.net/kb/> (version current as of October 20, 2006).

Tulving, E. and Thomson, D. (1973) Encoding Specificity and Retrieval Processes in Episodic Memory. *Psychological Review*, 80, 352-372. http://dx.doi.org/10.1037/h0020071

van Gog, T. and Sweller, J. (2015). Not New, but Nearly Forgotten: the Testing Effect Decreases or even Disappears as the Complexity of Learning Materials Increases. *Educational Psychology Review* 27: 247. https://doi.org/10.1007/s10648-015-9310-x

Wheeler, M.A., Ewers, M., and Buonanno, J. (2003). Different rates of forgetting following study versus test trials. *Memory*, 11, 571–580.

Wing, E.A., Marsh, E.J., and Cabeza R. (2013). Neural correlates of retrieval-based memory enhancement: An fMRI study of the testing effect. *Neuropsychologia*, 51(12), 2360-70.

Wooldridge, C. L., Bugg, J. M., McDaniel, M. A., and Liu, Y. (2014). The testing effect with authentic educational materials: A cautionary note. *Journal of Applied Research in Memory and Cognition*, 3, 214–221.

APPENDIX A Regression Analysis Results (RA)-Final Exam

Since column width was limited on the tables the following abbreviations were used:

Dep. Var. – Dependent Variables

Sig. – Significance

Adj. – Adjusted

Fixed Factor Used

P3 – Final exam

Covariates Used

PreCut – Pretreatment Exam1 with the mean of 93 subtracted

PreCutSqrd – PreCut squared

PreCutCubed – PreCut cubed

C0 T1 – Treatment Category

C0 – Control

C1 – Treatment

StudyTm – The self-reported amount of study per week averaged for the semester

Attend. – Attendance for the semester

Interact1 – Interaction 1; Precut times the treatment variable

Interact2 – Interaction 2; Precut Squared times the treatment variable

Interact3 – Interaction 3; Precut cubed times the treatment variable

FA – Principal Component (Factor Analysis)

Analyses Used

Anal 1- Analysis 1; which included PreCut, AvgStudyTm and Interact1

Anal 2- Analysis 2; which included PreCut, AvgStudyTm without Interact1

Anal 3- Analysis 3; which included PreCut, without AvgStudyTm, with Interact1

Anal 4- Analysis 4; which included PreCut, without AvgStudyTm, without Interact1

Anal 5- Analysis 5; which included PreCutSqrd, PreCut, AvgStudyTm, Interact1, and Interact2

Anal 6a- Analysis 6a; which included PreCutSqrd, PreCut, AvgStudyTm, Interact1, and without Interact2

Anal 6b- Analysis 6b; which included PreCutSqrd, PreCut, AvgStudyTm, without Interact1, and with Interact2

Anal 6c- Analysis 6c; which included PreCutSqrd, PreCut, AvgStudyTm, without Interact1, and without Interact2

Anal 7- Analysis 7; which included PreCutSqrd, PreCut, without AvgStudyTm, with Interact1, and with Interact2

Anal 8- Analysis 8; which included PreCutSqrd, PreCut, without AvgStudyTm, without Interact1, and without Interact2

Anal 9- Analysis 9; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, Interact1, Interact2, and Interact3

Anal 10a- Analysis 10a; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, Interact1, Interact2, and without Interact3

Anal 10b- Analysis 10b; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, Interact1, without Interact2, and with Interact3

Anal 10c- Analysis 10c; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, without Interact1, with Interact2, and with Interact3

Anal 10d- Analysis 10d; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, Interact1, without Interact2, and without Interact3

Anal 10e- Analysis 10e; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, without Interact1, without Interact2, and with Interact3

Anal 10f- Analysis 10f; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, without Interact1, with Interact2, and without Interact3

Anal 10g- Analysis 10g; which included PreCutCubed, PreCutSqrd, PreCut, AvgStudyTm, without Interact1, without Interact2, and without Interact3

Anal 11- Analysis 11; which included PreCutCubed, PreCutSqrd, PreCut, without AvgStudyTm, with Interact1, Interact2, and Interact3

Anal 12- Analysis 12; which included PreCutCubed, PreCutSqrd, PreCut, without AvgStudyTm, without Interact1, without Interact2, and without Interact3

The same preceding analyses were used for each of the eight largest components: FA2, FA3, FA5, FA10, FA11, FA12, FA22 and FA25.

Dep. Var. and Anal#	General Linear Model: Univariate	Fixed Factor & Sig.	Covariates	Sig.	R^2	Adj. R^2	F Statistic
P3 Anal 1	PreCut with StudyTm & Interact1 (PreCut * Treatment)	C0 T1 .541	1.PreCut 2.StdyTm 3.Interact1	.047 .344 .327	.302	.206	$F(4,29) = 3.135$ $p= .029$*
P3 Anal 2	PreCut with StudyTm	C0 T1 .558	1.PreCut 2.StdyTm NO Interact1	.044 .358 X	.278	.206	$F(3,30) = 3.850$ $p= .019$*
P3 Anal 3	PreCut with Interact1	C0 T1 .396	1.PreCut NO StdyTm 2.Interact1	.032 X .340	.280	.208	$F(3,30) = 3.881$ $p = .019$*

P3 Anal 4	PreCut	C0 T1 .413	1.PreCut NO StdyTm NO Interact1	.019 X X	.257	.209	F(2,31) =5.360 p=.010**
P3 Anal 5	PreCutSqr & Interact2 (PreCut^2 * Treatment); PreCut & Interact1; StdyTm	C0 T1 .604	1.PreCutSqr 2.PreCut 3.StdyTm 4.Interact1 5.Interact2	.591 .991 .303 .602 .752	.322	.172	F(6,27) = 2.139 p = .081
P3 Anal 6a	PreCutSqr; PreCut & Interact1; StdyTm	C0 T1 .364	1.PreCutSqr 2.PreCut 3.StdyTm 4.Interact1 NO Interact2	.4 00 .447 .300 .671 X	.320	.198	F(5,28) = 2.631 p = .045*
P3 Anal 6b	PreCutSqr & Interact2; PreCut; StdyTm	C0 T1 .527	1.PreCutSqr 2.PreCut 3.StdyTm NO Interact1 4.Interact2	.853 .367 .318 X .961	.315	.193	F(5,28) = 2.578 p = .049*
P3 Anal 6c	PreCutSqr; PreCut; StdyTm	C0 T1 .412	1.PreCutSqr 2.PreCut 3.StdyTm NO Interact1 NO Interact2	.219 .020 .310 X X	.315	.221	F(4,29) = 3.336 p= . 023*
P3 Anal 7	PreCutSqr & Interact2; PreCut & Interact1	C0 T1 .505	1.PreCutSqr 2.PreCut NO StdyTm 3.Interact1 4.Interact2	.638 925 X .674 .777	.295	.169	F(5,28) = 2.338 p = .068
P3 Anal 8	PreCutSqr PreCut	C0 T1 .292	1.PreCutSqr 2.PreCut NO StdyTm NO Interact1 NO Interact2	.247 .010 X X X	.290	.219	F(3,30) =4.804 p=.015**
P3 Anal 9	PreCutCube & Interact3 (PreCut^3 * Treatment); PreCutSqr & Interact2; PreCut & Interact1; StdyTm	C0 T1 .679	1.PreCutCub e 2.PreCutSqr 3.PreCut 4.StdyTm 5.Interact1 6.Interact2 7.Interact3	.853 .792 .856 .292 .626 .906 .812	.326	.111	F(8,25) = 1.515 p= .202

P3 Anal 10A	PreCutCube; PreCutSqr & Interact2; PreCut & Interact1; StdyTm	C0 T1 .551	PreCutCube PreCutSqr PreCut StdyTm Interact1 Interact2 NO Interact3	.751 .884 .916 .290 .557 .842 X	.325	.143	F(7,26) = 1.787 p= .133
P3 Anal 10B	PreCutCube & Interact3; PreCutSqr; PreCut & Interact1; StdyTm	C0 T1 .473	PreCutCube PreCutSqr PreCut StdyTm Interact1 NO Interact2 Interact3	.821 .613 .844 .280 .579 X .771	.326	.145	F(7,26) = 1.797 p= .131
P3 Anal 10C	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut; StdyTm	C0 T1 .571	PreCutCube PreCutSqr PreCut StdyTm NO Interact1 Interact2 Interact3	.716 .775 .618 .320 X .784 .692	.320	.137	F(7,26) = 1.747 p= .142
P3 Anal 10D	PreCutCube & Interact3; PreCutSqr; PreCut; StdyTm	C0 T1 .499	PreCutCube PreCutSqr PreCut StdyTm Interact1 NO Interact2 NO Interact3	.686 .469 .980 .289 .565 X X	.324	.174	F(6,27) = 2.155 p= .079
P3 Anal 10E	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut & Interact1; StdyTm	C0 T1 .581	PreCutCube PreCutSqr PreCut StdyTm NO Interact1 NO Interact2 Interact3	.766 .923 .402 .329 X X .755	.318	.166	F(6,27) = 2.097 p = .087
P3 Anal 10F	PreCutCube; PreCutSqr & Interact2; PreCut; StdyTm	C0 T1 .663	PreCutCube PreCutSqr PreCut StdyTm NO Interact1 Interact2 NO Interact3	.895 .849 .751 .340 X .910 X	.316	.164	F(6,27) = 2.076 p= .090

P3 Anal 10G	PreCutCube; PreCutSqr; PreCut; StdyTm	C0 T1 .464	PreCutCube PreCutSqr PreCut StdyTm NO Interact1 NO Interact2 NO Interact3	.933 .530 .104 .319 X X X	.315	.193	F(5,28) = .5792 p= .049*
P3 Anal 11	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut & Interact1	C0 T1 .690	PreCutCube PreCutSqr PreCut NO StdyTm Interact1 Interact2 Interact3	.888 .835 .921 X .779 .873 .884	.295	.105	F(7,26) = 1.555 p= .193
P3 Anal 12	PreCutCube; PreCutSqr; PreCut	C0 T1 .358	PreCutCube PreCutSqr PreCut NO StdyTm NO Interact1 NO Interact2 NO Interact3	.940 .553 .076 X X X X	.290	.192	F(4,29) = 2.963 p= .036*

APPENDIX B CA and RA Results FA Final Exam

FA = Component from Factor Analysis using the Principal Component Method

PreCut = Exam1-mean (93)

Treatment1 Control0; Attend. = Attendance

Component	Cronbach's Alpha	N
Col1_FA	.861	16
Col2_FA	.857	12
Col3_FA	.865	12
Col5_FA	.799	12
Col10_FA	.787	11
Col11_FA	.738	10
Col12_FA	.662	8
Col22_FA	.787	9
Col25_FA	.853	11

Dep Var Anal#	Gen. Linear Model: Univariate	Fixed Factor(s) & Sig.	Covariate(s)	Sig.	R^2	Adj. R^2	F Statistic
• FA1-1 / P3 Anal 1	PreCut with Interaction1 (PreCut * Treatment) & Attendance	Trtmnt1 Cntrl0 .381 / Trtmnt1 Cntrl0 .541	1.PreCut 2.Interact1 3.Attend.	.561 .418 .768	.088	-.038	F(4,29)= .601 p=.088 / F(4,29) = 3.135 p= .029
FA1-2 / P3 Anal 2	PreCut with Attendance	Trtmnt1 Cntrl0 .366 / Trtmnt1 Cntrl0 .558	1.PreCut No Interact1 2.Attend.	.895 X .752	.066	-.027	F(3,30)= .553 p=.066 / F(3,30)= 3.850 p=.019
FA1-3 / P3 Anal 3	PreCut with Interaction1	Trtmnt1 Cntrl0 .370 / Trtmnt1 Cntrl0 .396	1.PreCut 2.Interact1 No Attend.	.552 .407 X	.085	.007	F(3,30)= .927 p=.440 / F(3,30) = 3.881 p = .019

FA1-4 P3 Anal 4	PreCut	Trtmnt1 Cntrl0 .365 Trtmnt1 Cntrl0 .413	1.PreCut No Interact1 No Attend.	.892 X X	.257	.209	F(2,31)= 1.047 p=.363 F(2,31) =5.360 p=.010
FA1-5 P3 Anal 5	PreCutSqr & Interact2 (PreCut² * Treatment); PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .997 Trtmnt1 Cntrl0 .604	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 5.Attend.	.807 .801 .981 .965 .748	.107	-.091	F(6,27)= .542 p=.772 F(6,27) =2.139 p=.081
FA1-6a P3 Anal 6a	PreCutSqr; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 . 685 Trtmnt1 Cntrl0 .364	1.PreCutSqr 2.PreCut 3.Interact1 No Interact2 4.Attend.	.438 .420 .671 X .739	.107	-.052	F(5,28)= .674 p=-.647
FA1-6b P3 Anal 6b	PreCutSqr & Interact2; PreCut; Attend.	Trtmnt1 Cntrl0 .697 Trtmnt1 Cntrl0 .527	1.PreCutSqr 2.PreCut No Interact1 3.Interact2 4.Attend.	.464 .447 X .670 .740	.107	-.052	F(5,28)= .674 p=.647
FA1-6c P3 Anal 6c	PreCutSqr; PreCut; Attend.	Trtmnt1 Cntrl0 . 522 Trtmnt1 Cntrl0 .412	1.PreCutSqr 2.PreCut No Interact1 No Interact2 3.Attend.	.296 .302 X X .753	.102	-.022	F(4,29)= .819 p=.524
FA1-7 P3 Anal 7	PreCutSqr & Interact2; PreCut & Interact1	Trtmnt1 Cntrl0 .967 Trtmnt1 Cntrl0 .505	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 No Attend.	.780 774 .951 .936 X	.104	-.056	F(5,28)= .650 p=.664
FA1-8 P3 Anal 8	PreCutSqr PreCut	Trtmnt1 Cntrl0 . 524 Trtmnt1 Cntrl0 .292	1.PreCutSqr 2.PreCut No Interact1 No Interact2 No Attend.	.288 .294 X X X	.098	.008	F(3,30)= 1.091 p=.368

FA2-1 P3 Anal 1	PreCut with Interaction1 & Attend.	Trtmnt1 Cntrl0 .164 Trtmnt1 Cntrl0 .541	1.PreCut 2.Interact1 3.Attend.	.106 .225 .501	.496	.426	F(4,29)= 7.131 p=.000
FA2-2 P3 Anal 2	PreCut with Attend.	Trtmnt1 Cntrl0 .023 ** Trtmnt1 Cntrl0 .558	1.PreCut No Interact1 2.Attend.	.000 X .524	.469	.416	F(3,30)= 8.838 p=.000 F(3,30)= 3.850 P=019
FA2-3 P3 Anal 3	PreCut with Interaction1	Trtmnt1 Cntrl0 .164 Trtmnt1 Cntrl0 .396	1.PreCut 2.Interact1 No Attend.	.100 .226 X	.488	.437	F(3,30)= 9.524 p=.000
FA2-4 P3 Anal 4	PreCut	Trtmnt1 Cntrl0 .020 ** Trtmnt1 Cntrl0 .413	1.PreCut No Interact1 No Attend.	.000 X X	.462	.427	F(2,31)= 13.299 p=.000 F(2,31)= 5.360 p=010
FA2-5 P3 Anal 5	PreCutSqr & Interact2; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .590 Trtmnt1 Cntrl0 .604	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 5.Attend.	.776 .757 .595 .605 .578	.513	.405	F(6,27)= 4.740 p=.002
FA2-6a P3 Anal 6a	PreCutSqr; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .799 Trtmnt1 Cntrl0 .364	1.PreCutSqr 2.PreCut 3.Interact1 No Interact2 4.Attend.	.412 .467 .754 X .532	.508	.420	F(5,28)= 5.783 p=.001
FA2-6b P3 Anal 6b	Precut Cubed PreCutSqr & Interact2; PreCut; Attend.	Trtmnt1 Cntrl0 .860 Trtmnt1 Cntrl0 .527	1.PreCutSqr 2.PreCut No Interact1 3.Interact2 4.Attend.	.466 .515 X .774 .530	.508	.420	F(5,28)= 5.777 p=.001

FA2-6c P3 Anal 6c	PreCutCube PreCutSqr; PreCut; Attend.	Trtmnt1 Cntrl0 .052 * Trtmnt1 Cntrl0 .412	1.PreCutSqr 2.PreCut No Interact1 No Interact2 3.Attend.	.150 .278 X X .513	.506	.438	F(4,29)= 7.435 p=.000
FA2-7 P3 Anal 7	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut & Interact1	Trtmnt1 Cntrl0 .539 Trtmnt1 Cntrl0 .505	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 No Attend.	.733 .714 .546 .556 X	.507	.419	F(5,28)= 5.765 p=.001
FA2-8 P3 Anal 8	PreCutSqr PreCut	Trtmnt1 Cntrl0 .047 ** Trtmnt1 Cntrl0 .292	1.PreCutSqr 2.PreCut No Interact1 No Interact2 No Attend.	.147 247 X X X	.499	.449	F(3,30)= 9.953 p=.000 F(3,30)= 4.804 p=.015
FA3-1 P3 Anal 1	PreCut with Interaction1 & Attend.	Trtmnt1 Cntrl0 .294 Trtmnt1 Cntrl0 .541	1.PreCut 2.Interact1 3.Attend.	.062 .260 .085	.217	.109	F(4,29)= 2.008 p=.120
FA3-2 P3 Anal 2	PreCut with Attend.	Trtmnt1 Cntrl0 .347 Trtmnt1 Cntrl0 .558	1.PreCut No Interact1 2.Attend.	.096 X .081	.181	.099	F(3,30)= 2.215 p=.107
FA3-3 P3 Anal 3	PreCut with Interaction1	Trtmnt1 Cntrl0 .294 Trtmnt1 Cntrl0 .396	1.PreCut 2.Interact1 No Attend.	.067 .259 X	.131	.044	F(3,30)= 1.504 p=.234
FA3-4 P3 Anal 4	PreCut	Trtmnt1 Cntrl0 .328 Trtmnt1 Cntrl0 .413	1.PreCut No Interact1 No Attend.	.110 X X	.092	.034	F(2,31)= 1.577 p=.223

FA3-5 P3 Anal 5	PreCutSqr & Interact2; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .157 Trtmnt1 Cntrl0 .604	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 5.Attend.	.286 .299 .164 .171 .054	.296	.139	F(6,27)= 1.889 p=.119
FA3-6a P3 Anal 6a	PreCutSqr; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .593 Trtmnt1 Cntrl0 .364	1.PreCutSqr 2.PreCut 3.Interact1 No Interact2 4.Attend.	.326 .277 .618 X .077	.244	.109	F(5,28)= 1.807 p=.144
FA3-6b P3 Anal 6b	PreCutCube PreCutSqr & Interact2; PreCut; Attend.	Trtmnt1 Cntrl0 .615 Trtmnt1 Cntrl0 .527	1.PreCutSqr 2.PreCut No Interact1 3.Interact2 4.Attend.	.390 .341 X .665 .078	.242	.107	F(5,28)= 1.791 p=.147
FA3-6c P3 Anal 6c	PreCutCube PreCutSqr; PreCut; Attend.	Trtmnt1 Cntrl0 .209 Trtmnt1 Cntrl0 .412	1.PreCutSqr 2.PreCut No Interact1 No Interact2 3.Attend.	.156 .122 X X .077	.237	.132	F(4,29)= 2.253 p=.088
FA3-7 P3 Anal 7	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut & Interact1	Trtmnt1 Cntrl0 .251 Trtmnt1 Cntrl0 .505	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 No Attend.	.396 .411 .259 .266 X	.190	.046	F(5,28)= 1.315 p=.286
FA3-8 P3 Anal 8	PreCutSqr PreCut	Trtmnt1 Cntrl0 .200 Trtmnt1 Cntrl0 .292	1.PreCutSqr 2.PreCut No Interact1 No Interact2 No Attend.	.170 135 X X X	.149	.063	F(3,30)= 1.745 p=.179
FA5-1 P3 Anal 1	PreCut with Interaction1 & Attend.	Trtmnt1 Cntrl0 .340 Trtmnt1 Cntrl0 .541	1.PreCut 2.Interact1 3.Attend.	.012 .265 .460	.283	.184	F(4,29)= 2.855 p=.041

FA5-2 P3 Anal 2	PreCut with Attend.	Trtmnt1 Cntrl0 .047 ** Trtmnt1 Cntrl0 .558	1.PreCut No Interact1 2.Attend.	.005 X .446	.251	.176	F(3,30)= 3.343 p=.032 F(3,30)= 3.850 p=.019
FA5-3 P3 Anal 3	PreCut with Interaction1	Trtmnt1 Cntrl0 .329 Trtmnt1 Cntrl0 .396	1.PreCut 2.Interact1 No Attend.	.011 .253 X	.269	.196	F(3,30)= 3.673 p=.023
FA5-4 P3 Anal 4	PreCut	Trtmnt1 Cntrl0 .042 ** Trtmnt1 Cntrl0 .413	1.PreCut No Interact1 No Attend.	.005 X X	.236	.186	F(2,31)= 4.777 p=.016 F(2,31)= 5.360 p=.010
FA5-5 P3 Anal 5	PreCutSqr & Interact2; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .920 Trtmnt1 Cntrl0 .604	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 5.Attend.	.698 .669 .965 .999 .417	.337	.189	F(6,27)= 2.283 p=.065
FA5-6a P3 Anal 6a	PreCutSqr; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .305 Trtmnt1 Cntrl0 .364	1.PreCutSqr 2.PreCut 3.Interact1 No Interact2 4.Attend.	.142 .108 .339 X .406	.337	.218	F(5,28)= 2.841 p=.034
FA5-6b P3 Anal 6b	PreCutCube PreCutSqr & Interact2; PreCut; Attend.	Trtmnt1 Cntrl0 .278 Trtmnt1 Cntrl0 .527	1.PreCutSqr 2.PreCut No Interact1 3.Interact2 4.Attend.	.160 .126 X .340 .409	.337	.218	F(5,28)= 2.841 p=.034
FA5-6c P3 Anal 6c	PreCutCube PreCutSqr; PreCut; Attend.	Trtmnt1 Cntrl0 .020 ** Trtmnt1 Cntrl0 .412	1.PreCutSqr 2.PreCut No Interact1 No Interact2 3.Attend.	.112 .070 X X .436	.314	.220	F(4,29)= 3.321 p=.023

FA5-7 P3 Anal 7	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut & Interact1	Trtmnt1 Cntrl0 .995 Trtmnt1 Cntrl0 .505	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 No Attend.	.640 .612 .961 .927 X	.320	.198	F(5,28)= 2.634 p=.045
FA5-8 P3 Anal 8	PreCutSqr PreCut	Trtmnt1 Cntrl0 .018 ** Trtmnt1 Cntrl0 .292	1.PreCutSqr 2.PreCut No Interact1 No Interact2 No Attend.	.109 .068 X X X	.299	.229	F(3,30)= 4.274 p=.013 F(3,30)= 4.804 p=.015
FA10-1 P3 Anal 1	PreCut with Interaction1 & Attend.	Trtmnt1 Cntrl0 .222 Trtmnt1 Cntrl0 .541	1.PreCut 2.Interact1 3.Attend.	.028 .212 .613	.326	.233	F(4,29)= 3.502 p=.019
FA10-2 P3 Anal 2	PreCut with Attend.	Trtmnt1 Cntrl0 .783 Trtmnt1 Cntrl0 .558	1.PreCut No Interact1 2.Attend.	.038 X .595	.288	.217	F(3,30)= 4.042 p=.016
FA10-3 P3 Anal 3	PreCut with Interaction1	Trtmnt1 Cntrl0 .212 Trtmnt1 Cntrl0 .396	1.PreCut 2.Interact1 No Attend.	.025 .202 X	.320	.252	F(3,30)= 4.697 p=.008
FA10-4 P3 Anal 4	PreCut	Trtmnt1 Cntrl0 .763 Trtmnt1 Cntrl0 .413	1.PreCut No Interact1 No Attend.	.036 X X	.281	.235	F(2,31)= 6.059 p=.006
FA10-5 P3 Anal 5	PreCutSqr & Interact2; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .206 Trtmnt1 Cntrl0 .604	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 5.Attend.	.262 .276 .204 .202 .507	.370	.230	F(6,27)= 2.642 p=.038

FA10 - 6a P3 Anal 6a	PreCutSqr; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .958 Trtmnt1 Cntrl0 .364	1.PreCutSqr 2.PreCut 3.Interact1 No Interact2 4.Attend.	.673 .584 .945 X .603	.330	.210	F(5,28)= 2.759 p=.038
FA10 - 6b P3 Anal 6b	PreCutCube PreCutSqr & Interact2; PreCut; Attend.	Trtmnt1 Cntrl0 .923 Trtmnt1 Cntrl0 .527	1.PreCutSqr 2.PreCut No Interact1 3.Interact2 4.Attend.	.766 .683 X .897 .605	.330	.211	F(5,28)= 2.762 p=.038
FA10 - 6c P3 Anal 6c	Precut Cube PreCutSqr; PreCut; Attend.	Trtmnt1 Cntrl0 .561 Trtmnt1 Cntrl0 .412	1.PreCutSqr 2.PreCut No Interact1 No Interact2 3.Attend.	.188 .139 X X .593	.330	.237	F(4,29)= 3.570 p=.017
FA10 - 7 P3 Anal 7	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut & Interact1	Trtmnt1 Cntrl0 .226 Trtmnt1 Cntrl0 .505	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 No Attend.	.281 .296 .223 .220 X	.359	.245	F(5,28)= 3.141 p=.022
FA10 - 8 P3 Anal 8	PreCutSqr PreCut	Trtmnt1 Cntrl0 .541 Trtmnt1 Cntrl0 .292	1.PreCutSqr 2.PreCut No Interact1 No Interact2 No Attend.	.182 134 X X X	.323	.255	F(3,30)= 4.775 p=.008
FA11 - 1 P3 Anal 1	PreCut with Interaction1 & Attend.	Trtmnt1 Cntrl0 .378 Trtmnt1 Cntrl0 .541	1.PreCut 2.Interact1 3.Attend.	.909 .462 .850	.175	.061	F(4,29)= 1.537 p=.218
FA11 - 2 P3 Anal 2	PreCut with Attend.	Trtmnt1 Cntrl0 .061* Trtmnt1 Cntrl0 .558	1.PreCut No Interact1 2.Attend.	.407 X .835	.159	.075	F(3,30)= 1.891 p=.152

FA11 - 3 P3 Anal 3	PreCut with Interaction1	Trtmnt1 Cntrl0 .367 Trtmnt1 Cntrl0 .396	1.PreCut 2.Interact1 No Attend.	.905 .451 X	.174	.091	F(3,30)= 2.104 p=.121
FA11 - 4 P3 Anal 4	PreCut	Trtmnt1 Cntrl0 .058 * Trtmnt1 Cntrl0 .413	1.PreCut No Interact1 No Attend.	.399 X X	.158	.103	F(2,31)= 2.905 p=.070
FA11 - 5 P3 Anal 5	PreCutSqr & Interact2; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .251 Trtmnt1 Cntrl0 .604	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 5.Attend.	.578 .577 .289 .325 .997	.266	.103	F(6,27)= 1.630 p=.177
FA11 - 6a P3 Anal 6a	PreCutSqr; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .089 * Trtmnt1 Cntrl0 .364	1.PreCutSqr 2.PreCut 3.Interact1 No Interact2 4.Attend.	.137 .139 .100 X .910	.239	.103	F(5,28)= 1.754 p=.155
FA11 - 6b P3 Anal 6b	PreCutCube PreCutSqr & Interact2; PreCut; Attend.	Trtmnt1 Cntrl0 .091* Trtmnt1 Cntrl0 .527	1.PreCutSqr 2.PreCut No Interact1 3.Interact2 4.Attend.	.140 .141 X .110 .901	.234	.097	F(5,28)= 1.711 p=.165
FA11 - 6c P3 Anal 6c	PreCutCube PreCutSqr; PreCut; Attend.	Trtmnt1 Cntrl0 .078 * Trtmnt1 Cntrl0 .412	1.PreCutSqr 2.PreCut No Interact1 No Interact2 3.Attend.	.892 .944 X X .838	.160	.044	F(4,29)= 3.377 p=.266
FA11 - 7 P3 Anal 7	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut & Interact1	Trtmnt1 Cntrl0 . 239 Trtmnt1 Cntrl0 .505	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 No Attend.	.569 .568 .277 .313 X	.266	.135	F(5,28)= 2.028 p=.105

FA11 - 8 P3 Anal 8	PreCutSqr PreCut	Trtmnt1 Cntrl0 .074 * Trtmnt1 Cntrl0 .292	1.PreCutSqr 2.PreCut No Interact1 No Interact2 No Attend.	.170 .135 X X X	.158	.074	F(3,30)= 1.882 p=.154
FA12 - 1 P3 Anal 1	PreCut with Interaction1 & Attend.	Trtmnt1 Cntrl0 .202 Trtmnt1 Cntrl0 .541	1.PreCut 2.Interact1 3.Attend.	.016 .158 .218	.244	.140	F(4,29)= 2.343 p=.078
FA12 - 2 P3 Anal 2	PreCut with Attend.	Trtmnt1 Cntrl0 .088 * Trtmnt1 Cntrl0 .558	1.PreCut No Interact1 2.Attend.	.027 X .213	.189	.108	F(3,30)= 2.337 p=.094
FA12 - 3 P3 Anal 3	PreCut with Interaction1	Trtmnt1 Cntrl0 . 197 Trtmnt1 Cntrl0 .396	1.PreCut 2.Interact1 No Attend.	.016 .153 X	.203	.123	F(3,30)= 2.546 p=.075
FA12 - 4 P3 Anal 4	PreCut	Trtmnt1 Cntrl0 .082 * Trtmnt1 Cntrl0 .413	1.PreCut No Interact1 No Attend.	.029 X X	.146	.091	F(2,31)= 2.643 p=.087
FA12 - 5 P3 Anal 5	PreCutSqr & Interact2; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .227 Trtmnt1 Cntrl0 .604	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 5.Attend.	.188 .201 .212 .198 .180	.292	.135	F(6,27)= 1.857 p=.125
FA12 - 6a P3 Anal 6a	PreCutSqr; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .465 Trtmnt1 Cntrl0 .364	1.PreCutSqr 2.PreCut 3.Interact1 No Interact2 4.Attend.	.772 .882 .430 X .232	.247	.112	F(5,28)= 1.833 p=.139

FA12 - 6b P3 Anal 6b	PreCutCube PreCutSqr & Interact2; PreCut; Attend.	Trtmnt1 Cntrl0 .460 Trtmnt1 Cntrl0 .527	1.PreCutSqr 2.PreCut No Interact1 3.Interact2 4.Attend.	.672 .763 X .396 .230	.249	.115	F(5,28)= .859 p=.134
FA12 - 6c P3 Anal 6c	PreCutCube PreCutSqr; PreCut; Attend.	Trtmnt1 Cntrl0 .054 * Trtmnt1 Cntrl0 .412	1.PreCutSqr 2.PreCut No Interact1 No Interact2 3.Attend.	.230 .170 X X .211	.229	.123	F(4,29)= 2.157 p=.099
FA12 - 7 P3 Anal 7	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut & Interact1	Trtmnt1 Cntrl0 .294 Trtmnt1 Cntrl0 .505	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 No Attend.	.236 .251 .275 .257 X	.243	.107	F(5,28)= 1.793 p=.147
FA12 - 8 P3 Anal 8	PreCutSqr PreCut	Trtmnt1 Cntrl0 .050 ** Trtmnt1 Cntrl0 .292	1.PreCutSqr 2.PreCut No Interact1 No Interact2 No Attend.	.233 .173 X X X	.186	.104	F(3,30)= 2.283 p=.099 F(3,30)= 4.804 p=.015
FA22 - 1 P3 Anal 1	PreCut with Interaction1 & Attend.	Trtmnt1 Cntrl0 .336 Trtmnt1 Cntrl0 .541	1.PreCut 2.Interact1 3.Attend.	.861 .420 .127	.243	.139	F(4,29)= 2.330 p=.080
FA22 - 2 P3 Anal 2	PreCut with Attend.	Trtmnt1 Cntrl0 .047 ** Trtmnt1 Cntrl0 .558	1.PreCut No Interact1 2.Attend.	.410 X .120	.226	.148	F(3,30)= 2.916 p=.050 F(3,30)= 3.850 p=.019
FA22 - 3 P3 Anal 3	PreCut with Interaction1	Trtmnt1 Cntrl0 . 331 Trtmnt1 Cntrl0 .396	1.PreCut 2.Interact1 No Attend.	.845 .409 X	.179	.097	F(3,30)= 2.179 p=.111

FA22 - 4 P3 Anal 4	PreCut	Trtmnt1 Cntrl0 .059 * Trtmnt1 Cntrl0 .413	1.PreCut No Interact1 No Attend.	.416 X X	.160	.106	F(2,31)= 2.947 p=.067
FA22 - 5 P3 Anal 5	PreCutSqr & Interact2; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .546 Trtmnt1 Cntrl0 .604	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 5.Attend.	.888 .886 .597 .642 .163	.286	.127	F(6,27)= 1.801 p=.137
FA22 - 6a P3 Anal 6a	PreCutSqr; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .152 Trtmnt1 Cntrl0 .364	1.PreCutSqr 2.PreCut 3.Interact1 No Interact2 4.Attend.	.242 .247 .168 X .141	.280	.151	F(5,28)= 2.178 p=.085
FA22 - 6b P3 Anal 6b	PreCutCube PreCutSqr & Interact2; PreCut; Attend.	Trtmnt1 Cntrl0 .146 Trtmnt1 Cntrl0 .527	1.PreCutSqr 2.PreCut No Interact1 3.Interact2 4.Attend.	.235 .237 X .175 .139	.278	.149	F(5,28)= 2.159 p=.088
FA22 - 6c P3 Anal 6c	PreCutCube PreCutSqr; PreCut; Attend.	Trtmnt1 Cntrl0 .066 * Trtmnt1 Cntrl0 .412	1.PreCutSqr 2.PreCut No Interact1 No Interact2 3.Attend.	.753 .804 X X .126	.228	.122	F(4,29)= 2.146 p=.100
FA22 - 7 P3 Anal 7	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut & Interact1	Trtmnt1 Cntrl0 .450 Trtmnt1 Cntrl0 .505	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 No Attend.	.790 .787 .498 .541 X	.231	.094	F(5,28)= .686 p=.171
FA22 - 8 P3 Anal 8	PreCutSqr PreCut	Trtmnt1 Cntrl0 .081 * Trtmnt1 Cntrl0 .292	1.PreCutSqr 2.PreCut No Interact1 No Interact2 No Attend.	.755 .805 X X X	.163	.079	F(3,30)= 1.940 p=.144

FA25 - 1 P3 Anal 1	PreCut with Interaction1 & Attend.	Trtmnt1 Cntrl0 .455 Trtmnt1 Cntrl0 .541	1.PreCut 2.Interact1 3.Attend.	.505 .480 .041	.180	.067	F(4,29)= .591 p=.203
FA25 - 2 P3 Anal 2	PreCut with Attend.	Trtmnt1 Cntrl0 .574 Trtmnt1 Cntrl0 .558	1.PreCut No Interact1 2.Attend.	.870 X .037	.166	.082	F(3,30)= 1.983 p=.138
FA25 - 3 P3 Anal 3	PreCut with Interaction1	Trtmnt1 Cntrl0 .454 Trtmnt1 Cntrl0 .396	1.PreCut 2.Interact1 No Attend.	.507 .474 X	.050	-.045	F(3,30)= .528 p=.667
FA25 - 4 P3 Anal 4	PreCut	Trtmnt1 Cntrl0 .652 Trtmnt1 Cntrl0 .413	1.PreCut No Interact1 No Attend.	.886 X X	.033	-.029	F(2,31)= .537 p=.590
FA25 - 5 P3 Anal 5	PreCutSqr & Interact2; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .649 Trtmnt1 Cntrl0 .604	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 5.Attend.	.891 .899 .670 .691 .038	.211	.035	F(6,27)= 1.202 p=.335
FA25 - 6a P3 Anal 6a	PreCutSqr; PreCut & Interact1; Attend.	Trtmnt1 Cntrl0 .542 Trtmnt1 Cntrl0 .364	1.PreCutSqr 2.PreCut 3.Interact1 No Interact2 4.Attend.	.346 .328 .536 X .037	.206	.064	F(5,28)= 1.453 p=.236
FA25 - 6b P3 Anal 6b	PreCutCube PreCutSqr & Interact2; PreCut; Attend.	Trtmnt1 Cntrl0 .561 Trtmnt1 Cntrl0 .527	1.PreCutSqr 2.PreCut No Interact1 3.Interact2 4.Attend.	.376 .359 X .549 .038	.205	.063	F(5,28)= 1.447 p=.238

FA25 - 6c P3 Anal 6c	PreCutCube PreCutSqr; PreCut; Attend.	Trtmnt1 Cntrl0 .756 Trtmnt1 Cntrl0 .412	1.PreCutSqr 2.PreCut No Interact1 No Interact2 3.Attend.	.312 .307 X X .038	.195	. 084	F(4,29)= 1.756 p=.165
FA25 - 7 P3 Anal 7	PreCutCube & Interact3; PreCutSqr & Interact2; PreCut & Interact1	Trtmnt1 Cntrl0 .844 Trtmnt1 Cntrl0 .505	1.PreCutSqr 2.PreCut 3.Interact1 4.Interact2 No Attend.	.954 .946 .861 .878 X	.071	-.095	F(5,28)= .429 p=.824
FA25 - 8 P3 Anal 8	PreCutSqr PreCut	Trtmnt1 Cntrl0 .832 Trtmnt1 Cntrl0 .292	1.PreCutSqr 2.PreCut No Interact1 No Interact2 No Attend.	.336 .332 X X X	.063	-.030	F(3,30)= .676 p=.574

∗ 0.1 level

∗∗0.05 level

• Each individual factor (which consisted of a grouping of eight to sixteen final exam questions) was examined for similarities between the questions, to include: tissue types, necessity of recognition of a cell's internal characteristics, shapes of cells, orientation of micrographs, contrasts as a result of staining, and question “linguistic” variations such as question length and/or complexity. The twelve values greater than .3 were reviewed by an experienced professor and it was determined that Column 2 PCA represents an integration of staining and functional recognition and Column 5 PCA represents tissue organization and unique features. These were the only two columns that showed a consistent significance.

Appendix C Eigenvalues from Factor Analysis

Eigenvalues greater than 1 - Components 1 through 26; Total Variance Explained							
Compone nt	Initial Eigenvalu es			Extracti on Sums of Squared Loading s			Rotation Sums of Squared Loading sa
	Total	% of Varian ce	Cumulati ve %	Total	% of Varian ce	Cumulati ve %	Total
1	15.983	14.53	14.53	15.983	14.53	14.53	9.358
2	9.736	8.851	23.381	9.736	8.851	23.381	7.698
3	7.977	7.252	30.633	7.977	7.252	30.633	7.181
4	7.617	6.925	37.558	7.617	6.925	37.558	7.67
5	5.675	5.159	42.717	5.675	5.159	42.717	4.836
6	5.301	4.819	47.536	5.301	4.819	47.536	4.808
7	5.074	4.613	52.149	5.074	4.613	52.149	4.237
8	4.827	4.388	56.536	4.827	4.388	56.536	4.937
9	4.419	4.018	60.554	4.419	4.018	60.554	5.281
10	4.099	3.727	64.281	4.099	3.727	64.281	6.5
11	3.661	3.328	67.609	3.661	3.328	67.609	5.697
12	3.471	3.156	70.764	3.471	3.156	70.764	3.661
13	3.294	2.995	73.759	3.294	2.995	73.759	3.701
14	2.949	2.681	76.44	2.949	2.681	76.44	4.767
15	2.808	2.553	78.993	2.808	2.553	78.993	3.265
16	2.52	2.291	81.284	2.52	2.291	81.284	4.17
17	2.31	2.1	83.384	2.31	2.1	83.384	3.403
18	2.242	2.038	85.423	2.242	2.038	85.423	3.529
19	2.117	1.925	87.348	2.117	1.925	87.348	4.939
20	1.98	1.8	89.148	1.98	1.8	89.148	4.646
21	1.68	1.528	90.675	1.68	1.528	90.675	3.615
22	1.499	1.363	92.038	1.499	1.363	92.038	4.086
23	1.324	1.204	93.241	1.324	1.204	93.241	3.975
24	1.278	1.162	94.403	1.278	1.162	94.403	5.286

	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
25	1.104	1.004	95.407	1.104	1.004	95.407	6.466
26	1.002	0.911	96.318	1.002	0.911	96.318	5.07
27	0.822	0.748	97.066				
28	0.807	0.734	97.799				
29	0.703	0.639	98.438				
30	0.554	0.504	98.942				
31	0.477	0.434	99.376				
32	0.419	0.381	99.757				
33	0.268	0.243	100				

Extraction Method: Principal Component Analysis.
a When components are correlated,
sums of squared loadings cannot be added to obtain a total variance.

Appendix D Pattern Matrix

	1	2	3	5	10	11	12	22	25
1a								.417	
1b						.336			
2a				.432				.376	
2b									
3a									
3b				.338					
4a						.962			
4b						.400			
6a			.362						
6b									
7b		.352							
9a									
9b	.340	.470							
10a			.679						
10b		.353	.427						.330
11a									.443

11b									
12b									
13a			.956						
13b			.956						
14a					.406				
14b									
15a				.520	.367				
15b	.355								
16a	.373					.340			
16b									
17							.950		
18a			.353		.754				
18b					.413		.386		
19a					.317				
19b							.351		
19c		.320							
19d					.349				
20a	.308								.427
20b	.320							.347	.310
20c	.328								
21				.309					
22a									.381
22b		-.329						.431	.320

22c									.688
23a	.385		.537						
23b				.831					
24a	.341	.302					.345		
24b									
25a						.778			
25b									
26a									
26b									
27b								.966	
27c			.497		.437				
28a									
28b									
29a			.956						
29b					.868				
30b		.301	.410						
31a						.314			
31b			.337						
32a									
32b									
32c		.574					.372		
32d									
33a	.671								

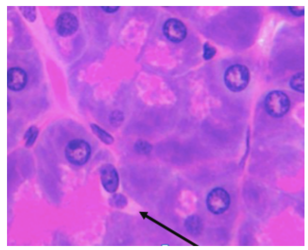
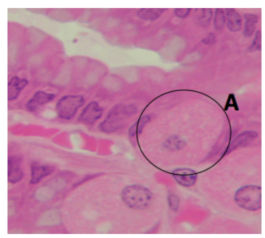
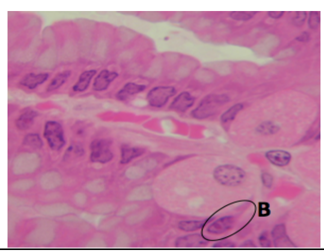
33b									.309
34a						.452			-.352
34b									
35a									
35b									
35c		.776							
36a	.359			.639					
36b				.384					
37b									
38a									.664
38b						.816			
39b									
40a									.788
40b									.778
41b				.952					
42a	.446								
42b									
43a	.909								
43b		.424		.394					
44a		.791							
44b		.691							
45a									
45b									

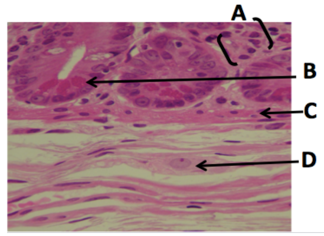
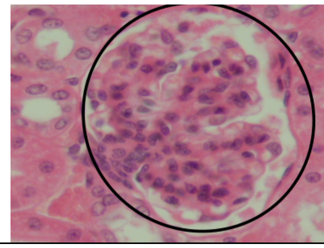
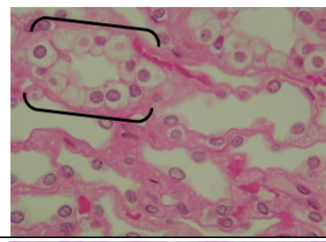
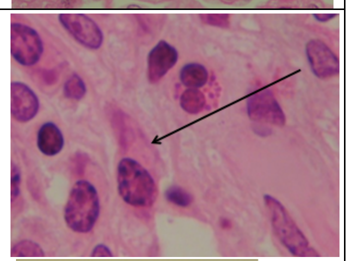
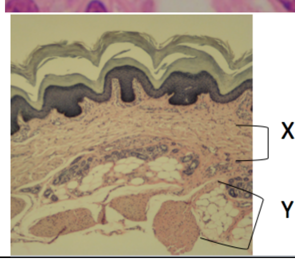
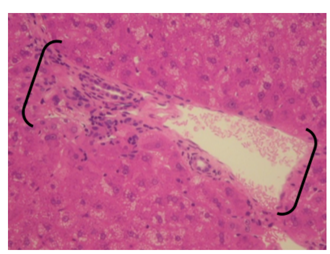
46a				.430			.317		
46b				.494					
47a									
47b				.314					
47c						.512			
47d									
48a					.366			.305	
48b					.346		.312	.366	
49									
50a					.311				
50b								.326	
51									
52a									
52b									
53a									
53b			.361						
54a	.969								
54b	.671								
56	.671								
57a		.567							
57b									
58								.371	
59a						.433			

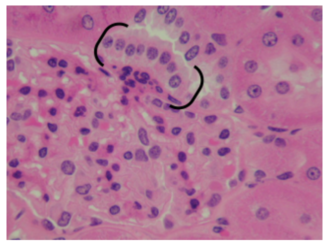
59b							.352		
60	.446								

Appendix E PCA Column 2

PCA Column 2 Integration of staining and functional recognition

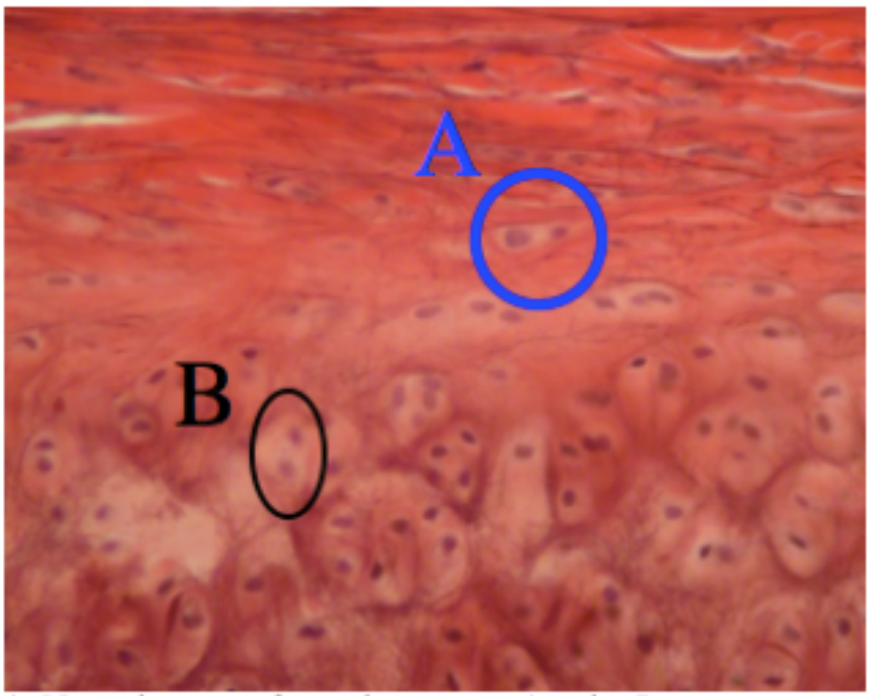
Pattern Matrixa			
Component			
	2	2	
7b	0.352	Secretion of bicarbonate to protect duct from active enzymes & increase pH of duodenum (since stomach is pH2) or beginning of duct	
9b	0.470	HCl; Intrinsic Factor	
10b	0.353	Pepsinogen	

19c	0.320	Smooth muscle, muscularis mucosa	
24a	0.302	fenestrated epithelium	
30b	0.301	Renal medulla (has thin loop of Henle)	
32c	0.574	C.Clock-face nucleus	
35c	0.776	c. Hypodermis	
43b	0.424	Sinusoids, full credit; Central Vein/partial credit	

44a	0.791	primary	
44b	0.691	myeloperoxidase	Same as above
57a	0.567	Macula densa	
		4	
Extraction Method: Principal Component Analysis.			
a. Rotation converged in 63 iterations.			

Appendix F Sample LE question and feedback

Question 1 (0.05 points)
LE 8A A



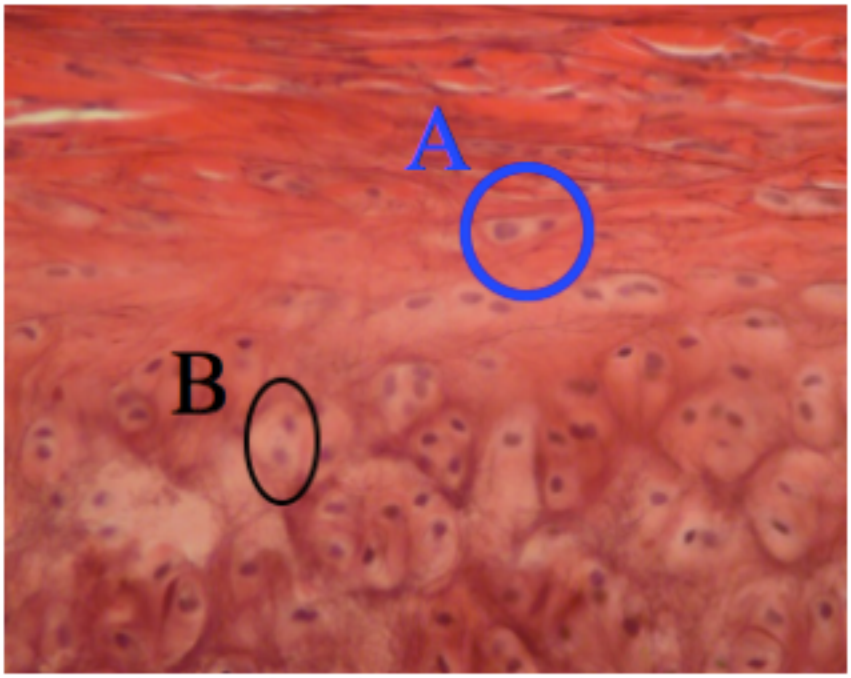
1. Name the types of growth present at A and at B.
Type your answer to A in the 1ST box.
Type your answer to B in the 2ND box.
Type "X" in the 3rd box to make this easier to give attendance credit.

Question 1 options:

Answer Blank # 1	
Answer Blank # 2	
Answer Blank # 3	

Save

Question 2 Feedback (0.05 points)
LE 8A A



1. Name the types of growth present at A and at B.
A. Appositional growth
B. Interstitial growth
With respect to A there is a perichondrium adjacent to the cartilage. The type of cartilage growth that is exhibited within the Dense Irregular CT is appositional growth (this growth continues throughout adulthood, grows only at edges, and widens cartilage). With respect to B, the cartilage is elastic cartilage. The type of growth that is exhibited within the elastic cartilage (in fetal tissue and young children) is interstitial growth and can grow in any direction.

Please place an "x" in the answer box to acknowledge that you reviewed this answer.

Question 3 options:

Save